

PHENET
PHENOTYPING & ENVIROTypING
SOLUTIONS FOR AGROECOLOGY

TOOLS AND METHODS FOR EXTENDED PLANT PHENOTYPING AND ENVIROTypING SERVICES OF EUROPEAN RESEARCH INFRASTRUCTURES

Deliverable D2.3

Computer vision models based on AI trained models

Due date (as in DoA):	June 2026
Actual achievement date:	June 2026
Work package:	2
Work package leader:	David Rousseau (UA)
Deliverable leader:	David Rousseau (UA)
Version:	1

TABLE OF CONTENT

1	Summary	4
2	Results	4
3	Partners involved in the work	4

1 SUMMARY

Based on the identified phenotyping traits to be automated through computer vision via devices designed within PHENET in deliverables D2.1 and D2.2 the partners involved in WP2 have developed computer vision models. These are presented in this document. The models are all made accessible with a standalone ergonomic and user friendly interface. Companion items such as research article, conference proceedings, video demo or data challenge may also be associated. As a demonstration of the progress toward Deliverable D2.4 (light version of these models embedded on devices) we introduce already some generic methodologies to design light models.

2 RESULTS

FrontVeg is a computer vision model designed for foreground vegetation segmentation in close-range images, with a specific focus on trellised crop systems (such as the one targeted in UC Orchard of PHENET). The model operates on RGB images acquired under field conditions. A specificity is that it does not need 3D LIDAR information to do this nor Flash light. Its processing pipeline consists of two main steps. First, a depth map is estimated from the RGB input and analysed using histogram-based thresholding to separate foreground and background regions. Second, a Grounding DINO model is applied to detect foliage within the extracted foreground. The method is implemented as a plugin for Napari, allowing both automated inference and optional manual refinement of the segmentation mask by the user. The output is a binary mask where foreground vegetation is highlighted. Beyond the sole apple orchard use, the model has been successfully applied to trellised crops such as apricot, and grapevine.

The model is freely accessible as a Napari model : <https://napari-hub.org/plugins/frontveg.html>

Companion research article (published) : Metuarea, Herearii, et al. "Individual Segmentation of Intertwined Apple Trees in a Row via Prompt Engineering." *Sensors* 25.15 (2025): 4721

<https://www.mdpi.com/1424-8220/25/15/4721>

Companion video presented during the Society of European Botany conference 2025 by Herearii Metuarea : https://youtu.be/cvW3jmiQprY?si=L1xu6U5tFr_I4L_j

DeepPhenoTree is a computer vision model designed for the automatic detection of key phenological stages in perennial crops, focusing on flowering, fruitlet development, and fruit at harvest time. The model operates on RGB images acquired in field conditions and targets structured orchard environments, particularly apple production systems. It relies on deep learning–based object detection using a RT-DETR architecture trained on the multi-site **DeepPhenoTree – Apple Edition** dataset, which includes annotated images from several European countries. The model performs targeted inference for three specific tasks—flowering (BBCH 00–69), fruitlet (BBCH 71–77), and fruit (BBCH 81–89)—and outputs bounding boxes corresponding to detected organs, along with object counts. It is implemented as a plugin for napari, allowing users to run inference interactively on their own images or on provided sample datasets. Performance evaluation based on cross-validation across four sites shows robust results, with mean average precision (mAP@0.5) reaching approximately 0.65 for flowering, 0.82 for fruitlet, and 0.86 for fruit detection, with small variability among the 5 sites of the orchard site of the UC orchards.

The model is freely accessible as a Napari model : <https://napari-hub.org/plugins/deepphenotree.html>

Companion data paper article (under review) : Metuarea, Herearii, et al. "DeepPhenoTree–Apple Edition: a Multi-site apple phenology RGB annotated dataset with deep learning baseline models." (2026). <https://www.researchsquare.com/article/rs-8977752/v1>

Companion Data challenge organized in the framework of the ECCV 2026 Computer Vision for Plant Phenotyping and Agriculture Workshop : <https://cvppa2026.github.io/challenges>

LightPhenosuite : This suite of deep learning models include of light version of standard models in plant phenotyping specifically useful for the UC GxE wheat such as the Vegann model (<https://github.com/simonMadec/VegAnn>) segmenting the soil covering stages or the wheat ear detection from the Global wheat data challenge (<https://www.global-wheat.com/>). The version of these models proposed in the LightPhenosuite are made of machine learning–based computer vision at the edge for intelligent plant imaging, with the objective of enabling real-time monitoring of plant health and growth on embedded devices. The approach targets resource-constrained hardware platforms such as single-board computers, smartphones, and microcontrollers, and addresses the technical challenges associated with deploying deep learning models in such environments. The methodology focuses on evaluating model optimisation strategies, including quantisation, pruning, knowledge distillation, and the use of lightweight architectures of the type Kolmogorov Arnold Networks, in order to reduce computational cost while maintaining acceptable performance.

The models and the methodology on how to produce new light versions are accessible at : <https://forge.inrae.fr/imhorphen/LightPhenosuite>

Companion book chapter (published) : Garnier, Julien, et al. "Edge-Based Machine Learning for Computer Vision in Smart Plant Biology Imaging." *Machine Learning for Plant Biology* (2025): 1-12.

Companion video presented during the 2025 International Plant Phenotyping Symposium of the IPPN by Julien Garnier : <https://youtu.be/N9Wa8K9H9vY?si=WU0GVc6-UbqB5coR>

MPM (Mutual Pair Merging) is a computer vision method designed to improve the efficiency of Vision Transformer (ViT) models for semantic segmentation by reducing the number of processed tokens during inference. The approach operates on image tokens extracted from RGB images and introduces a training-free token aggregation mechanism that merges similar tokens based on cosine similarity. Specifically, tokens are paired using a mutual nearest-neighbor criterion and averaged, producing a reduced token set while preserving essential information. A key feature of the method is the use of a merge map that enables reconstruction of the original token grid before the segmentation decoder, ensuring compatibility with standard architectures such as Segmenter without requiring architectural modifications. The model does not rely on learned parameters or continuous compression thresholds; instead, the trade-off between computational efficiency and accuracy is controlled through the selection of insertion points within the transformer encoder. Experimental results reported in the article show that MPM can significantly reduce inference latency—up to 60% on edge devices and notable throughput gains on high-end GPUs—while maintaining limited accuracy degradation (typically below 3% mIoU). The method is particularly effective in deployment scenarios requiring real-time or resource-constrained processing, although it may introduce slight smoothing effects on fine structures due to token aggregation and incurs a quadratic cost for similarity computation at very high resolutions. As such MPM could be applied to all developed models provided that they are based on vision transformers (these being the current state of the art in terms of computer vision at that time).

MPM'ed transformers for the UC GxE Wheat of PHENET are accessible at : <https://forge.inrae.fr/imhorphen/mpm>

Companion conference proceedings (Accepted at CVPR 2026) : Ravé, Simon et al. "MPM: Mutual Pair Merging for Efficient Vision Transformers." *arXiv preprint arXiv:2604.05718* (2026). <https://arxiv.org/pdf/2604.05718>

Code is available at: <https://forge.inrae.fr/imhorphen/mpm>

COYL (Counting Orange and Yellow Larvae) is an application dedicated to count automatically on RGB images, the orange wheat blossom midge, *Sitodiplosis mosellana*, and the yellow wheat blossom midge, *Contarinia tritici* which are major pests of wheat crops. This study originated from a practical problem encountered by the breeders: the need to increase the throughput of larvae counting to characterize the susceptibility of numerous wheat varieties. This problem was solved using a deep learning method applied to RGB images acquired using smartphones. Several YOLO-based object detection architectures were evaluated through a coarse-to-fine strategy to identify the most effective configuration. The best-performing model (YOLOv26m) achieved high detection performance. The model successfully distinguished between the two visually similar species, although slightly better performance was observed for *S. mosellana*. The output for the breeders was an online application to count the larvae. It necessitated inputs from agronomists, data scientists, and web-development specialists. At the time of writing these lines, the online application is only available for the members of the Walloon Agricultural Research Centre, but further web developments could be made to secure the connection and open it to the web.

The complete process, from the problem to the practical solution, including all stages of research is described in

Companion journal article (under submission at Plant Phenomics Journal):

Antoine Deryck et al. "Automatic detection and counting of *Contarinia tritici* and *Sitodiplosis mosellana* larvae using smartphone images and deep learning"

A concrete output for the community is the release of the labeled dataset. This dataset will be shortly online. It has been named the COYL-1 (Counting Orange and Yellow Larvae) dataset. It can be used to train wheat blossom midge detection models, or it can be included in a more generic dataset, for example to count larvae for any insect species. Final model and code will be available at the same location.

DOI of the application : <https://doi.org/10.5281/zenodo.19402333>

iphone_orange_22.jpg



Nombre de larves total : 268

Nombre de larves oranges : 268

Nombre de larves jaunes : 0

© CRAW 2025 - v2.0 - Application de comptage de larves de cécidomyies

Développé par: Sébastien DANDRIFOSSE & Antoine DERYCK

Mis en ligne par: Patrick HOUBEN

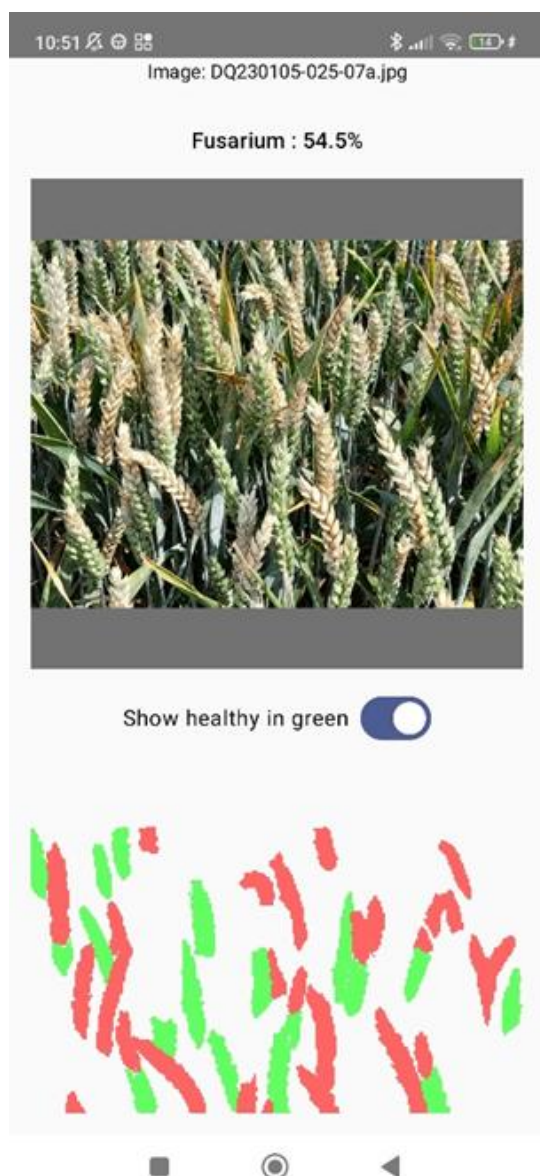
FUSASEYD (Assessment of FHB) is an application dedicated to the assessment, on RGB images, of ears showing FHB symptoms. The breeding process leading to the creation of disease-resistant cereals requires characterizing and assessing many genotypes. Up to now, only an expert eye can assess the disease susceptibility of cereal varieties in field trials. In this framework, developing automated assessment methods based on imaging techniques could provide useful information, helping the experts in their task of identifying disease-resistant varieties. This study highlights the potential of RGB imaging associated with an instance segmentation model based on Yolo model to detect and quantify FHB on winter wheat ears. This study led to a proof of concept for the use of RGB images and deep learning instead of visual counting to classify the varietal resistance against FHB for the purposes of plant breeding and registration in the catalog.

GEVES has developed a PC user interface and a smartphone application, called “FUSASEYD”, to visualize on the Web the symptoms predicted by the model. This application design still needs improvement and will be validated during the next campaign 2026 in the trials for registration in France.

For the next step, it is planned to contribute to the Global Wheat datasets (Hund, 2020) to improve precision phenotyping under field conditions by assembling large, diverse and well annotated image data and by updating models.

The full study is described in:

Companion journal article (under review at Journal of Experimental Botany: “Plant Phenomics & Enviromics Across Scales” special issue): V. Cadot et al. Resistance assessment of wheat varieties against fusarium head blight using artificial intelligence on RGB images.



Fusaseyd application to quantify Fusarium Head Blight

3 PARTNERS INVOLVED IN THE WORK

(UA) and all partners of the UC GxE Orchard for the models deployable for Apple tree phenotyping.
 (GEVES - CRAW) & all partners of the UC Disease for the model deployable for Wheat disease scoring.
 (UA) & all partners of the UC GxE Wheat for the light models deployable on connected stick.